



# Penghitung orang pada video CCTV menggunakan metode histogram of oriented gradient dan filter Kalman

## People counter on CCTV video using histogram of oriented gradient and Kalman filter methods

Faisal Dharma Adhinata<sup>1)</sup>, Muhammad Ikhsan<sup>1)</sup>, Wahyono<sup>2\*)</sup>

<sup>1)</sup>Magister Ilmu Komputer, Fakultas MIPA, Universitas Gadjah Mada  
Sekip Utara, Bulaksumur, Yogyakarta, Indonesia 55281

<sup>2)</sup>Departemen Ilmu Komputer dan Elektronika, Fakultas MIPA, Universitas Gadjah Mada  
Sekip Utara, Bulaksumur, Yogyakarta, Indonesia 55281

**Cara sitasi:** D. A. Faisal, M. Ikhsan, and W. Wahyono, "Penghitung orang pada video CCTV menggunakan metode histogram of oriented gradient dan filter Kalman," Jurnal Teknologi dan Sistem Komputer, vol. 8, no. 3, pp. 222-227, 2020. doi: [10.14710/jtsiskom.2020.13660](https://doi.org/10.14710/jtsiskom.2020.13660), [Online].

**Abstract** - CCTV cameras have an important function in the field of public service, especially for convenience. The objects recorded through CCTV cameras are processed into information to support service satisfaction in the community. This study uses the function of CCTV for people counting from objects recorded by a camera. Currently, the process of detecting and tracking people takes a long time to detect all frames. In this study, the frame selection into keyframes uses the mutual information entropy method. The keyframes processing uses the Histogram of Oriented Gradient (HOG) and Kalman filter methods. The proposed method results F1 value of 0.85, recall of 76 %, and precision of 97 % with winStride parameter (12,12), scale 1.05, and the distance of the human object to CCTV 4 meters.

**Keywords** – people detection; HOG; Kalman filter; mutual information entropy; people counter

**Abstrak** - Kamera CCTV mempunyai peran penting dalam bidang pelayanan publik, terutama dalam hal kenyamanan. Objek yang terekam kamera CCTV diolah menjadi sebuah informasi sebagai penunjang kepuasan layanan kepada masyarakat. Penelitian ini menggunakan fungsi CCTV untuk menghitung jumlah orang yang terekam kamera. Proses deteksi dan pelacakan orang membutuhkan waktu yang lama jika harus mendeteksi pada semua frame. Metode mutual information entropy digunakan untuk menyeleksi frame menjadi keyframe. Keyframe yang diperoleh selanjutnya diolah menggunakan metode histogram of oriented gradient (HOG) dan filter Kalman. Metode yang diusulkan memberikan hasil nilai F1 0,85, recall 76 %, dan presisi 97 % dengan parameter winStride (12,12) dan scale 1,05 serta jarak objek manusia terhadap CCTV 4 meter.

**Kata kunci** - deteksi orang; HOG; filter Kalman; mutual informasi entropi; penghitung orang

### I. PENDAHULUAN

Penggunaan teknologi semakin meluas hingga ke berbagai aspek kehidupan. Dalam bidang pelayanan publik, penggunaan teknologi kamera pengawas (CCTV) menjadi penunjang dalam memberikan kenyamanan bagi masyarakat. Di sisi lain, para pemilik pelayanan publik, seperti perkantoran, menyadari bahwa data video kamera CCTV bisa menunjang keberhasilan dalam melayani segala kebutuhan masyarakat. Data rekaman video CCTV dapat digunakan untuk menghitung jumlah orang yang berada di kawasan tersebut dan mengetahui jumlah orang yang berkunjung [1].

Kajian penghitungan orang dengan media video dari kamera CCTV dikembangkan melalui bidang keilmuan *computer vision*. Pustaka OpenCV sering digunakan untuk mendeteksi objek, termasuk objek manusia [2], [3]. Metode yang sering digunakan untuk deteksi objek manusia dalam sebuah frame video adalah Histogram of Oriented Gradient (HOG) dan Continuously Adaptive Meanshift (CAMshift). Kidane [4] melakukan perbandingan metode HOG dan CAMshift untuk deteksi orang yang menghasilkan metode HOG lebih baik daripada metode CAMshift. Kajian lain menggunakan metode HOG dan Support Vector Machine (SVM) untuk mendeteksi pejalan kaki yang berjalan di jalan raya [5]. Kombinasi metode HOG dan filter Kalman untuk deteksi dan pelacakan manusia pada video [6].

Kajian penghitungan jumlah orang menggunakan HOG juga dilakukan dalam [7]. HOG merupakan deskriptor efektif untuk menangkap informasi intensitas gradient pada objek manusia, termasuk deteksi tubuh, kepala, dan bahu. Metode HOG tersebut dipilih untuk mendeteksi pejalan kaki dan hasil deteksi diterapkan pada awal tahap dengan basis informasi heuristik [8]. Hasil deteksi digunakan pada tahap pelacakan. Filter Kalman ditambahkan untuk melakukan pelacakan

\*) Penulis korespondensi (Wahyono)  
Email: wahyo@ugm.ac.id

dengan mencocokkan fitur bentuk secara visual yang diusulkan dalam [3]. Hasil yang diperoleh dilakukan klasifikasi untuk menentukan objek yang terdeteksi dan pelacakan manusia atau bukan. SVM cocok untuk mengklasifikasikan objek manusia pada studi kasus deteksi orang [9].

Salah satu tantangan pemrosesan data video adalah kecepatan metode ekstraksi fitur dalam mengolah setiap frame. Pengolahan dengan menggunakan semua frame video memperlambat pemrosesan sistem. Data video terlebih dahulu dilakukan tahap prapemrosesan. Video tersusun dari frame-frame yang berurutan. Objek yang tidak bergerak atau frame yang tidak mengandung objek bergerak di dalamnya hanya dihitung satu frame agar pengolahan frame video efisien. Sistem ini memerlukan metode untuk memilih frame yang hanya mengandung objek bergerak, atau disebut dengan pemilihan keyframe. Pemilihan keyframe dilakukan untuk mengurangi jumlah data yang diperlukan untuk pengindeksan video dan mengelola informasi yang dihasilkan [10]. Ouyang dkk. [11] melakukan perbandingan metode seleksi keyframe menggunakan *mutual information entropy* dan memberikan hasil terbaik untuk seleksi keyframe.

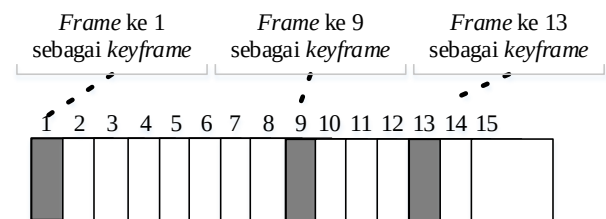
Kajian-kajian penghitungan objek manusia tersebut tidak mempertimbangkan kecepatan pemrosesan video secara *real-time*. Kecepatan pemrosesan *real-time* perlu diperbaiki karena metode HOG membutuhkan waktu yang lama dalam melakukan pengolahan setiap frame. Penelitian ini mengkaji aplikasi penghitungan orang yang terekam melalui kamera CCTV dengan menggunakan seleksi keyframe untuk meningkatkan kecepatan pemrosesan perhitungan jumlah orang secara *real-time*. Analisis penggunaan parameter HOG dilakukan untuk mencapai hasil deteksi orang yang optimal.

## II. METODE PENELITIAN

Penelitian ini memiliki langkah kerja yang terdiri dari akuisisi data, prapemrosesan, deteksi orang, pelacakan orang, perhitungan orang, dan evaluasi. Akuisisi data video dilakukan dengan perekaman kamera CCTV. Data video diekstrak menjadi frame-frame video. Frame-frame video diseleksi menjadi keyframe menggunakan metode *mutual information entropy*. Keyframe yang terpilih dilakukan deteksi objek manusia menggunakan metode HOG. Objek manusia ini dilakukan pelacakan menggunakan metode filter Kalman agar *bounding box* pada manusia tidak hilang pada keyframe selanjutnya. *Bounding box* yang melewati area tertentu digunakan untuk perhitungan jumlah orang.

### A. Akuisisi data

Proses akuisisi data video dalam penelitian ini menggunakan kamera CCTV dengan perekaman secara *real-time*. Video direkam dengan kamera CCTV *HiLook Wi-Fi PT Camera* dengan model IPC-P120-D/W. Resolusi kamera CCTV yang digunakan adalah 2.0 MP



Gambar 1. Proses seleksi keyframe

full HD (1920×1080) dengan perekaman 10 frame per detik (fps).

### B. Prapemrosesan data

Tahap prapemrosesan dilakukan untuk ekstraksi video menjadi frame. Frame-frame dilakukan seleksi keyframe dengan metode *mutual information entropy*. Nilai entropi menunjukkan keteracakan distribusi derajat keabuan suatu citra [12]. Semakin acak distribusi derajat keabuannya, semakin tinggi nilai entropi yang dihasilkan. *Mutual information entropy* menunjukkan dua frame (citra) yang memiliki informasi berbeda sehingga memerlukan perbandingan dari frame satu dengan frame lainnya [13].

Tahapan prapemrosesan berupa seleksi keyframe ditunjukkan pada Gambar 1. Frame pertama otomatis menjadi keyframe. Frame kedua dilakukan perhitungan *mutual information entropy* dengan keyframe terpilih (frame pertama). Jika nilainya kurang dari ambang batas, maka frame kedua disimpan sebagai keyframe baru. Jika nilai *mutual information entropy*-nya lebih dari ambang batas, maka frame tidak disimpan sebagai keyframe baru. Proses ini diulang sampai ditemukan keyframe baru, misalnya perbandingan keyframe 1 dengan frame ke-9, nilai *mutual information entropy*-nya kurang dari ambang batas, maka frame 9 disimpan sebagai keyframe baru. Proses ini akan diulang sampai frame terakhir hasil ekstraksi video.

### C. Deteksi Orang

Ekstraksi fitur pada kandidat objek menjadi kunci dalam melakukan deteksi. Dalam kajian ini, metode deteksi manusia dengan pendekatan klasifikasi menggunakan metode HOG dan SVM seperti dalam [14]. Gambar 2 menunjukkan tahapan metode HOG.

Ekstraksi fitur HOG diperoleh dari hasil perhitungan nilai *gradient* dalam daerah tertentu pada frame video. Tiap frame mempunyai karakteristik yang ditunjukkan oleh distribusi *gradient*. Karakteristik ini diperoleh dengan membagi frame ke dalam daerah kecil atau *cell*. Tiap *cell* disusun sebuah histogram dari sebuah *gradient*. Kombinasi dari histogram ini dijadikan sebagai deskriptor yang mewakili sebuah objek manusia [14].

Langkah pertama yang dilakukan adalah seleksi objek manusia dengan operasi HOG. Objek citra yang terdeteksi dihitung nilai deskriptornya menggunakan HOG. Perhitungan nilai *gradient* setiap piksel frame menggunakan Persamaan 1 dengan  $I$  adalah frame skala keabuan.  $I_x$  adalah matriks terhadap sumbu x dan  $I_y$

adalah matriks terhadap sumbu y. Matrik  $I_x$  dan  $I_y$  dihitung dengan Persamaan 2. Parameter  $D_x$  adalah mask  $[-1 \ 0 \ 1]$  dan  $D_y$  adalah mask  $[-1 \ 0 \ 1]^T$  yang masing-masing dihitung dengan cara konvolusi.

$$|G| = \sqrt{I_x^2 + I_y^2} \quad (1)$$

$$I_x = I * D_x, I_y = I * D_y \quad (2)$$

Dalam penghitungan histogram dan orientasi, hasil gradient ditransformasikan ke dalam koordinat sumbu dengan sudut antara 0 sampai  $180^\circ$  yang disebut orientasi gradient. Orientasi gradient ( $\theta$ ) dinyatakan dengan Persamaan 3. Setelah mendapat nilai *gradient*, maka proses selanjutnya adalah menentukan jumlah *bin* orientasi yang digunakan dalam histogram. Namun sebelum menghitung gradient, frame dibagi menjadi beberapa *cell* dan dikelompokkan menjadi ukuran lebih besar yang disebut *block*.

$$\theta = \tan^{-1}(I_x/I_y) \quad (3)$$

Pada proses normalisasi *block*, HOG memiliki jendela deteksi dengan ukuran  $64 \times 128$  yang terdiri dari  $8 \times 8$  piksel. Proses normalisasi *block* dengan pemisah beberapa elemen vektor ini menghasilkan fitur HOG [14]. Nilai normalisasi fitur *block* diperoleh dari Persamaan 4.

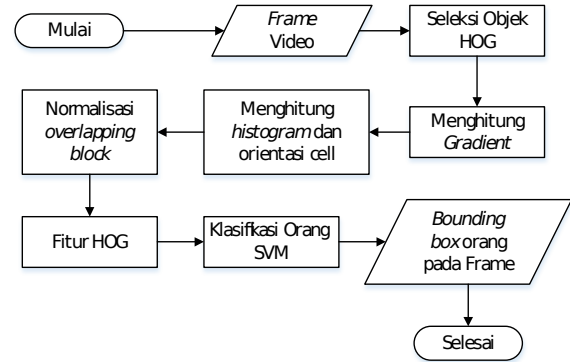
$$b = \frac{b}{\sqrt{b^2 + e}} \quad (4)$$

Fitur *block* dinormalisasi untuk mengurangi efek perubahan kecerahan objek pada satu *block*. Variabel  $b$  merupakan nilai blok fitur dan  $e$  merupakan bilangan bernilai kecil untuk mencegah pembagian dengan 0. Penentuan objek yang dideteksi orang atau bukan menggunakan SVM. Ide dasar dari SVM adalah memisahkan dua kelas melalui *hyperplane* yang ditentukan oleh vektor normal  $w$  dan bias  $b$ . *Hyperplane* pemisah yang optimal adalah yang memaksimalkan jarak antara *hyperplane* dan titik terdekat dari kedua kelas yang disebut margin [15].

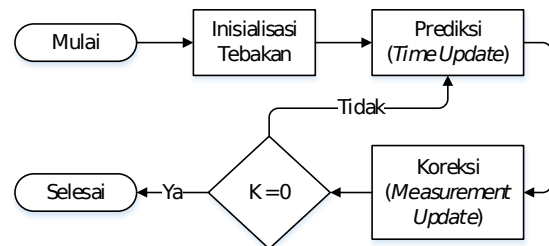
#### D. Pelacakan Orang

Objek yang terdeteksi dilakukan pelacakan dengan melihat perkembangan atau perubahan posisi dan kecepatannya menggunakan filter Kalman. Filter Kalman digunakan untuk mencari estimasi nilai variabel keadaan objek dalam proses yang sedang berlangsung dimana objek tersebut tidak dapat diprediksi secara acak, memiliki ketidakpastian dan variasi dengan model Linear Quadratic Estimation (LQE) [16].

Pada filter Kalman, estimasi dilakukan dengan dua tahapan, yaitu tahap prediksi (*time update*) dan tahap koreksi (*measurement update*) seperti ditunjukkan dalam Gambar 3. Tahap prediksi (*time update*) dilakukan untuk memproyeksikan nilai ke depan



Gambar 2. Tahapan metode HOG



Gambar 3. Tahapan pelacakan filter Kalman

keadaan saat ini dan perkiraan kesalahan *covariant* untuk mendapatkan apriori untuk waktu berikutnya [6]. Tahap koreksi (*measurement update*) melakukan respons atas hasil estimasi variabel keadaan. Variabel tersebut akan dikoreksi menggunakan model pengukuran. Variabel estimasi yang telah dikoreksi digunakan pada tahap prediksi (*time update*).

#### E. Skenario Pengujian

Pengujian sistem pada kajian ini dilakukan untuk penghitungan orang menggunakan metode HOG dan filter Kalman dan menerapkan seleksi *keyframe* untuk penyesuaian terhadap sumber daya komputasi. Pengujian meliputi pengujian waktu terhadap resolusi video, pengaruh seleksi *keyframe* terhadap waktu pemrosesan, pengujian parameter HOG, dan pengujian penghitungan orang. Pengujian resolusi video diperlukan untuk mengetahui pengaruh ukuran resolusi terhadap kecepatan. Pengujian untuk seleksi *keyframe* perlu dilakukan juga untuk mengetahui kecepatan seleksi *keyframe*. Pengujian seleksi *keyframe* dilakukan dengan video offline dengan durasi tertentu. Semua pengujian penelitian ini dilakukan di waktu siang hari.

Pengujian resolusi video dilakukan untuk membandingkan beberapa resolusi video. Tujuan pengujian ini untuk menentukan resolusi video yang memberikan waktu pemrosesan paling minimal. Pengujian dilakukan dengan resolusi video  $1920 \times 1080$ ,  $960 \times 540$ , dan  $640 \times 480$ . Pengujian seleksi *keyframe* dilakukan untuk membandingkan waktu pengolahan video dengan menggunakan seleksi *keyframe* dan tanpa seleksi *keyframe*. Pengujian parameter HOG dilakukan untuk membandingkan parameter *winStride* dan *scale* pada metode HOG dalam menentukan parameter yang

memberikan hasil optimal dalam deteksi objek. Kombinasi parameter *winStride* dan *scale* yang digunakan pada kajian ini adalah nilai *winStride* (4,4); (8,8); dan (12,12) yang dikombinasikan dengan parameter *scale* 1,05; 1,1; 1,2; dan 1,3. Pengujian perhitungan orang dilakukan untuk menghitung jumlah orang yang terdeteksi pada area yang telah ditentukan dalam video.

### III. HASIL DAN PEMBAHASAN

Penggunaan metode HOG dan filter Kalman untuk penghitungan jumlah orang dengan data video sangat lambat karena melakukan pemrosesan terhadap semua frame. Pemrosesan pada metode HOG tersebut mempunyai parameter *multiScale* yang membuat lambat karena dilakukan pada skala citra yang berbeda. Parameter tersebut mempunyai pengaruh terhadap deteksi *false* positif, yaitu deteksi pada lokasi tertentu yang sebenarnya tidak ada objek manusia namun terdeteksi sebagai objek manusia, dan terhadap kecepatan dalam mengolah frame video. Pengaruh tersebut perlu diperhatikan karena kajian ini menggunakan video secara *real-time*.

#### A. Pengujian resolusi video

Permasalahan kecepatan dalam mengolah frame video yang lambat dapat diselesaikan menggunakan *resize* resolusi frame video. Hasil pengujian pengaruh resolusi frame video terhadap kecepatan dinyatakan pada Tabel 1. Pengujian ini menghasilkan kecepatan pengolahan per frame paling baik adalah pada resolusi video 640×480 dengan waktu pengolahan 0,19 detik per frame. Pada resolusi 1920×1080 menghasilkan nilai fps 0 karena pengolahan per frame lebih dari 1 detik.

Pengolahan frame pada video dilakukan dengan mengolah piksel-piksel sesuai ukuran resolusi. Semakin kecil resolusi video yang dilakukan, maka semakin sedikit piksel yang diolah. Oleh karena itu, penurunan ukuran resolusi video menghasilkan penurunan waktu pemrosesan sesuai dengan [17]. Pemrosesan data video dilakukan dengan memproses citra berurutan. Namun, untuk mengurangi ukuran resolusi video harus memperhatikan detail objek yang diolah. Karena keterbatasan sumber daya memori, pengujian selanjutnya menggunakan resolusi 640×480.

#### B. Pengujian seleksi frame

Penggunaan seleksi keyframe bertujuan untuk menentukan frame yang diproses sehingga tidak semua frame digunakan. Hal ini bertujuan untuk mengurangi jumlah memori yang dibutuhkan dan waktu pemrosesan. Pengujian terhadap seleksi keyframe menggunakan video *offline* dengan durasi 17 detik dan jumlah frame 510 (30 fps).

Hasil pengujian pengaruh seleksi keyframe terhadap kecepatan ditunjukkan pada Tabel 2. Hasil seleksi keyframe pada video menghasilkan waktu pemrosesan video yang lebih cepat karena tidak semua frame diolah.

**Tabel 1.** Ukuran resolusi frame terhadap kecepatan

No	Resolusi video	Waktu proses setiap frame (detik)	Nilai fps
1	1920 × 1080	1,678	0
2	960 × 540	0,38	3
3	640 × 480	0,190	5

**Tabel 2.** Pengaruh seleksi keyframe terhadap kecepatan

No	Pengolahan	Waktu Pengolahan (detik)
1	Tanpa seleksi keyframe	199,975
2	Seleksi keyframe	44,569

Tahap proses seleksi frame sebagai keyframe membutuhkan waktu 6 mdetik untuk membandingkan dua frame. Hasil waktu pengolahan masih terdapat waktu tunda (seharusnya 17 detik). Namun, penggunaan seleksi keyframe ini cukup membantu dalam pengolahan *real-time* penghitungan jumlah orang yang terdeteksi di kamera. Hasil ini menunjukkan seleksi keyframe mampu mempercepat waktu proses karena jumlah data yang diproses jauh lebih sedikit seperti dinyatakan dalam [10].

#### C. Pengujian parameter HOG

Parameter HOG yang berpengaruh signifikan terhadap akurasi dan kecepatan adalah *winStride* dan *scale*. Parameter *winStride* merupakan ukuran *sliding window* yang digunakan untuk mendeteksi objek manusia. Ukuran parameter *winStride* yang kecil menyebabkan objek kecil dapat terdeteksi sebagai manusia sehingga banyak menimbulkan derau. Parameter *scale* digunakan sebagai jumlah level pada citra piramida. Ukuran parameter *scale* yang kecil menyebabkan meningkatnya jumlah lapis pada citra piramida sehingga waktu komputasinya semakin lama.

Pengujian pengaruh parameter *winStride* dan *scale* terhadap akurasi ditunjukkan pada Tabel 3. Pada percobaan ini menggunakan jumlah objek manusia satu orang dan peletakan kamera dengan tinggi 1 meter. Hasil tersebut menunjukkan terdapat tiga percobaan yang menghasilkan tingkat ketepatan deteksi lebih baik, yaitu percobaan ke 2, 3, dan 9.

Namun, dalam percobaan ke 2 dan 3 terdapat *false* positif, yaitu bukan manusia namun terdeteksi sebagai manusia oleh sistem. Dalam beberapa kali perekaman, objek manusia tidak terdeteksi oleh metode HOG. Contoh dalam Gambar 4 menunjukkan *false* negatif. Salah satu penyebab dari masalah ini adalah pose dari objek manusia yang dilihat dari samping padahal data latih yang dimiliki mayoritas adalah data yang diambil dari arah depan [6]. Dalam percobaan ke 7 dan 10, terdapat beberapa frame yang terdeteksi *false* negatif, yaitu terdapat objek manusia, namun tidak terdeteksi oleh sistem. Oleh karena itu, parameter *winStride* (12,12) dan *scale* 1,05 dipilih sebagai parameter dalam pengujian untuk perhitungan jumlah orang karena tidak



**Tabel 3.** Pengaruh parameter *winStride* dan *scale* terhadap *false positif* dan *true positif*

No	<i>winStride</i>	<i>Scale</i>	<i>False Positif</i>	<i>True Positif</i>
1	(4,4)	1,05	4	7
2	(4,4)	1,1	1	7
3	(4,4)	1,2	2	7
4	(4,4)	1,3	4	6
5	(8,8)	1,05	1	7
6	(8,8)	1,1	0	6
7	(8,8)	1,2	0	4
8	(8,8)	1,3	0	2
9	(12,12)	1,05	0	7
10	(12,12)	1,1	0	4
11	(12,12)	1,2	0	2
12	(12,12)	1,3	0	0

**Tabel 4.** Hasil pengaruh jarak terhadap nilai recall dan presisi

Jarak (m)	TP	FP	FN	Recall (%)	Presisi (%)	Skor F1
3	22	2	15	59	92	0,72
4	29	1	9	76	97	0,85

menimbulkan *false positif* dan objek manusia terdeteksi secara langsung.

#### D. Pengujian penghitungan orang

Pengujian penghitungan orang menggunakan parameter metode HOG dengan *winStride* (12,12), *scale* 1,05, dan tinggi kamera dengan lantai satu meter. Penghitungan orang dilakukan pada area tertentu, yaitu menggunakan area sebelah kanan kamera. Sebagai contoh, pengujian dilakukan dengan jarak kamera dengan objek manusia sekitar 3 meter dan 4 meter. Data latih yang digunakan metode HOG adalah keseluruhan badan manusia.

Hasil pengujian pengaruh jarak objek dengan kamera terhadap nilai recall dan presisi dinyatakan dalam Tabel 4. Objek yang terlalu dekat dengan kamera tidak bisa terdeteksi. Pada pengujian pada jarak 1 meter dan 2 meter, objek manusia tidak terdeteksi sehingga nilai *false negatif* tinggi. Jarak 1 dan 2 meter tidak dapat merekam objek manusia keseluruhan badan sehingga tidak terdeteksi oleh sistem. Hal ini karena masih banyaknya nilai *false negatif*, yaitu seharusnya terdeteksi ada orang, namun oleh sistem tidak terdeteksi.

Contoh *false negatif* pada pengujian penghitungan orang ditunjukkan dalam Gambar 5. Objek manusia pada region yang ditentukan harusnya berjumlah 3 orang, namun hanya terdeteksi 2 orang. Jika jarak perekaman antara objek dengan kamera lebih dari 4 meter, objek manusia kemungkinan masih terdeteksi karena objek manusia terlihat keseluruhan badannya. Namun, penelitian selanjutnya perlu dilakukan untuk membuktikannya. Contoh hasil *true positif* ditunjukkan dalam Gambar 6. Pengujian tersebut menunjukkan bahwa terdapat 3 orang yang berada dalam area sebelah



**Gambar 4.** Contoh *false negatif* pada pengujian parameter metode HOG



**Gambar 5.** Contoh *false negatif* pada pengujian penghitungan orang



**Gambar 6.** Contoh *true positif* pada pengujian penghitungan orang

kanan. Jumlah orang yang terdeteksi pada percobaan ini adalah 3.

Metode filter Kalman dalam kajian ini berpengaruh terhadap hasil deteksi orang. Hasil ini sesuai dengan [16] yang menyatakan bahwa metode filter menghasilkan ketepatan pelacakan pada pejalan kaki dari berbagai arah objek manusia. Penggunaan filter Kalman membuat objek manusia hanya terdeteksi satu orang sehingga tampak hasil *bounding box* tiap orang hanya satu sebagaimana ditampilkan pada Gambar 6.

Walaupun hasil yang diperoleh cukup baik, metode yang diusulkan memiliki kelemahan, yaitu masih dipengaruhi oleh hasil deteksi orang berbasis HOG+SVM yang masih terkendala oleh pose *invariant* dari manusia. Oleh sebab itu, pada penelitian selanjutnya disarankan untuk menggunakan berbagai macam pose dalam data

latih sehingga HOG+SVM dapat mendeteksi manusia dengan lebih baik.

#### IV. KESIMPULAN

Metode HOG dan filter Kalman telah diterapkan untuk penghitungan orang melalui kamera CCTV dengan menghasilkan nilai  $F1$  0,85, recall 76 %, dan presisi 97 % pada jarak 4 meter. Hasil akurasi dipengaruhi oleh jarak serta parameter *winStride* dan *scale*. Metode seleksi keyframe dapat mengurangi waktu pengolahan video real-time.

#### UCAPAN TERIMA KASIH

Peneliti berterimakasih kepada Lembaga Pengelola Dana Pendidikan (LPDP) yang telah memberikan beasiswa program Magister Indonesia berupa dana pendidikan dan penelitian ini.

#### DAFTAR PUSTAKA

- [1] D. Roqueiro and V. A. Petrushin, "Counting people using video cameras," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 22, no. 3, pp. 193-209, 2007. doi: [10.1080/17445760601139096](https://doi.org/10.1080/17445760601139096)
- [2] F. Sayadi, Y. Said, M. Atri, and R. Tourki, "Real time human detection in video streams," in *1st International Conference on Networks and Systems for Communication, Education and Data Processing*, Porto, Portugal, Jul. 2012, pp. 154-158.
- [3] C.-C. Chen, H.-H. Lin, and O. T.-C. Chen, "Tracking and counting people in visual surveillance systems," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Prague, Czech Republic, May 2011, pp. 1425-1428. doi: [10.1109/ICASSP.2011.5946681](https://doi.org/10.1109/ICASSP.2011.5946681)
- [4] H. L. Kidane, "Comparative survey: People detection, tracking and multi-sensor Fusion in video sequence," 2018. [Online]. Available: [arXiv:1806.06261](https://arxiv.org/abs/1806.06261)
- [5] L. Fei and L. D. Lin, "Pedestrian detection based on histogram of oriented gradient in intelligent transportation system," in *6th International Conference on Intelligent Networks and Intelligent Systems*, Shenyang, China, Nov. 2013, pp. 78-81. doi: [10.1109/ICINIS.2013.27](https://doi.org/10.1109/ICINIS.2013.27)
- [6] C. Li, L. Guo, and Y. Hu, "A new method combining HOG and Kalman filter for video-based human detection and tracking," in *3rd International Congress on Image and Signal Processing*, Yantai, China, Oct. 2010, pp. 290-293. doi: [10.1109/CISP.2010.5648239](https://doi.org/10.1109/CISP.2010.5648239)
- [7] J. Liu, J. Liu, and M. Zhang, "A detection and tracking based method for real-time people counting," in *2013 Chinese Automation Congress*, Hunan, China, Nov. 2013, pp. 470-473. doi: [10.1109/CAC.2013.6775780](https://doi.org/10.1109/CAC.2013.6775780)
- [8] X. Wang, C. Wang, and J. Yao, "A heuristic information based system for people counting," in *Third International Conference on Intelligent Human-Machine Systems and Cybernetics*, Hangzhou, China, Aug. 2011, pp. 22-26. doi: [10.1109/IHMSC.2011.12](https://doi.org/10.1109/IHMSC.2011.12)
- [9] F. Zhu, X. Yang, J. Gu, and R. Yang, "A new method for people-counting based on support vector machine," in *Second International Conference on Intelligent Networks and Intelligent Systems*, Tianjian, China, Nov. 2009, pp. 342-345. doi: [10.1109/ICINIS.2009.94](https://doi.org/10.1109/ICINIS.2009.94)
- [10] S. kaur Riya, "Review of purposed method for key frame extraction from videos," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 7, pp. 377-379, 2017. doi: [10.26483/ijarcs.v8i7.4279](https://doi.org/10.26483/ijarcs.v8i7.4279)
- [11] S. Z. Ouyang, L. Zhong, and R. Q. Luo, "The comparison and analysis of extracting video key frame," *IOP Conference Series: Materials Science and Engineering*, vol. 359, 2018. doi: [10.1088/1757-899X/359/1/012010](https://doi.org/10.1088/1757-899X/359/1/012010)
- [12] D. M. Gavrilu and S. Munder, "Multi-cue pedestrian detection and tracking from a moving vehicle," *International Journal of Computer Vision*, vol. 73, pp. 41-59, 2007. doi: [10.1007/s11263-006-9038-7](https://doi.org/10.1007/s11263-006-9038-7)
- [13] L. Sun and Y. Zhou, "A key frame extraction method based on mutual information and image entropy," in *International Conference on Multimedia Technology*, Hangzhou, China, Jul. 2011, pp. 35-38. doi: [10.1109/ICMT.2011.6001938](https://doi.org/10.1109/ICMT.2011.6001938)
- [14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, USA, Jun. 2005, pp. 886-893. doi: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177)
- [15] M. Pérez-Ortiz, S. Jiménez-Fernández, P. Gutiérrez, E. Alexandre, C. Hervás-Martínez, and S. Salcedo-Sanz, "A review of classification problems and algorithms in renewable energy applications," *Energies*, vol. 9, no. 8, 607, pp. 1-27, 2016. doi: [10.3390/en9080607](https://doi.org/10.3390/en9080607)
- [16] P. Chong and Y. H. Tay, "A novel pedestrian detection and tracking with boosted HOG classifiers and kalman filter," in *IEEE Student Conference on Research and Development*, Kuala Lumpur, Malaysia, Dec. 2016, pp. 1-5. doi: [10.1109/SCORED.2016.7810052](https://doi.org/10.1109/SCORED.2016.7810052)
- [17] K. Romic, I. Galic, and H. Leventic, "Influence of the input image resolution on the staircase detection," in *International Symposium ELMAR*, Zadar, Croatia, Sep. 2016, pp. 177-180. doi: [10.1109/ELMAR.2016.7731781](https://doi.org/10.1109/ELMAR.2016.7731781)